

NPS ARCHIVE
1964
THOMAS, E.

BETTER PERFORMANCE APPRAISAL
FOR A BETTER CIVIL SERVICE

EDWARD W. THOMAS

DUDLEY KNOX LIBRARY
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CA 93943-5401

Library
U. S. Naval Postgraduate School
Monterey, California

 John Smith

BETTER PERFORMANCE APPRAISAL
FOR A BETTER CIVIL SERVICE

* * * * *

Edward W. Thomas

BETTER PERFORMANCE APPRAISAL
FOR A BETTER CIVIL SERVICE

by

Edward W. Thomas
Commander, United States Navy

This document is subject to special export controls and each transmittal to foreign government or foreign nationals may be made only with prior approval of the U.S. Naval Postgraduate School (Code 035).

Submitted in partial fulfillment of
the requirements for the degree of

MASTER OF SCIENCE
IN
MANAGEMENT

United States Naval Postgraduate School
Monterey, California

1964

Library
U. S. Naval Postgraduate School
Monterey, California

BETTER PERFORMANCE APPRAISAL
FOR A BETTER CIVIL SERVICE

by

Edward W. Thomas

This work is accepted as fulfilling
the requirements for the degree of
MASTER OF SCIENCE

IN

MANAGEMENT

from the

United States Naval Postgraduate School

ABSTRACT

Properly applied, performance evaluation can be of value in promoting good teamwork, increasing work efficiency, and assisting in the development of effective managers. The study searches for methods of improving Civil Service performance appraisal systems within the constraints of presently existing laws and regulations. The evolution of present systems is traced, and the resulting practices are examined in the light of recent research and industrial experience. Alternative appraisal methods are discussed, and broad suggestions for improvement are offered.

TABLE OF CONTENTS

CHAPTER	PAGE
I. THE PROBLEM--DEVELOPMENT OF MANAGERS	1
Management Development in the Civil Service .	2
Methods to Improve	8
Performance Evaluation	10
Summary	12
II. PRESENT EVALUATION PROCEDURES	15
Performance Rating Act of 1950	17
Evaluation for Morale and Efficiency	24
Evaluation for Management Information	28
Second Hoover Commission	34
Summary	36
III. DISCUSSION OF ALTERNATIVES	38
Purposes of Evaluation	39
Evaluation for Promotion	44
Confidential Reports	48
Methods of Evaluation	52
Evaluating the Evaluator	58
Summary	61
IV. SUGGESTED IMPROVEMENTS	64
BIBLIOGRAPHY	72

CHAPTER I

THE PROBLEM--DEVELOPMENT OF MANAGERS

The qualities of the public service required by (the events of this generation) are almost self-evident. High competence in administration; integrity, stability and reliability in performance; and most significantly, the capacity for innovation and creativity are needed to envision and attain national goals.¹

Criticism of the public service has been a popular pastime in all countries and all ages. Even the best-run bureaucracies have inherent shortcomings which attract unfavorable attention while their usually unspectacular achievements go unnoticed. It is not surprising, therefore, that the American Civil Service in the middle years of the twentieth century should continue to bear the popular image of hopeless red tape, waste and inefficiency. Actually the several recent impartial studies of the subject suggest that the service as a whole forms a surprisingly effective instrument, considering its relatively unplanned and piece-meal historical development.

Today, 80 years after the Pendleton Act made the merit system a permanent fixture in our government, the Federal Civil Service numbers over two and one-half million people, which represents over three per cent of the country's entire labor force. Their salaries alone currently comprise

¹Wallace S. Sayre, "The Public Service", Goals for Americans, (New York: Prentice Hall, 1963), p. 285.

nearly one-eighth of the National budget. In view of this high cost in money and manpower and the important part played by the Civil Service in the success or failure of the American Nation, thoughtful citizens are rightly concerned with the economy and effectiveness of the service and its responsiveness to the people's best interests.

This is an age of accelerating technological advancement, and people have become, in an increasing degree, dependent on each other. Individuals now perform small parts of larger processes. The days of simple, relatively independent enterprises seems to be passing, and the operation of today's organizations involves the cooperative efforts of many component elements and different technical skills. Indeed the size and complexity of organizations today, both in and out of government, far surpass anything we have known before.

For all of these reasons, the direction and coordination of government undertakings challenge the wisdom and ingenuity of its managers as never before. Expanding technology, increased organizational size and sheer numbers of factors going into decisions now require management of the highest order.

MANAGEMENT DEVELOPMENT IN THE CIVIL SERVICE

The most important controllable factor in the

continued fruitful existence of any organization is the quality of its personnel--particularly of those occupying positions of leadership. If a group is to maintain its vitality and efficient productivity for any length of time, it must devote serious thought and effort to the development of competent administrators.

One might assume that competent management could be taken for granted in an organization so large and so vital to the country's welfare as the executive branch of the government. Unfortunately, too many people have taken it for granted. The Federal Civil Service on the whole has been slow to avail itself of the advantages to be gained through a systematic effort to recognize, nourish and exploit management potential in its employees. Writing as recently as 1952, one authority stated:

It is a disturbing fact that most departments and agencies and the Civil Service Commission have not developed adequate programs for promoting employees within the Federal Service. All too often advancements take place not because of any conscious or deliberate effort by responsible officials or agencies, and not because of any "plan" intentionally designed to help and guide the employees, or any "system" furnishing positive and constructive assistance to the persons involved. On the contrary, in most cases, progress is made without benefit of such a plan, and despite the system now in existence.²

²G. Lyle Belsley, Federal Personnel Management and the Transition, (Washington: Public Administration Clearing House, 1952) p. 38.

Quite recently a young naval officer, nearing the completion of a tour of duty in a large federal agency and liking the work very much, seriously considered resigning from the Navy and accepting a Civil Service position there. However, in deciding against such a course he stated, "The only trouble is that I would hate to have my livelihood dependent upon the personnel processes of this agency. They leave good supervisors in the same jobs year after year and promote all the wrong people. As a matter of fact, any correspondence between efficiency ratings and actual performance is pretty much coincidental."

Such criticisms, however, are not applicable to the Civil Service as a whole. The employees of the federal government comprise not one, but a fairly loose federation of services. Merit systems independent of the regular service, for instance, are employed by the Tennessee Valley Authority, Federal Bureau of Investigation, Atomic Energy Commission, the Foreign Service and others. Furthermore, the Civil Service Commission has increasingly decentralized administrative functions of the regular service to the different departments and agencies which now operate in accordance with fairly broad standards and guidelines provided by the Commission. Although the different agencies are in general still subject to similar laws and regulations,

the result has been a significant divergency in practice and performance among the various organizations.³

Although there remain serious shortcomings in the personnel practices of many units, the records of others have been outstandingly good. The Navy's supervisory selection program for blue-collar workers, for instance, has been singled out as being well conceived and adapted to the Navy's needs.⁴ The Tennessee Valley Authority's career system has also been cited as being a good example of what a federal personnel program ought to be, and reasonably satisfactory programs have been developed in some other departments.⁵

Yet, many agencies, some of them very active in the field of technical training, still fail to face squarely the needs for executive development. The factors contributing to slow progress are readily identifiable, in many cases stemming from the historical traditions of

³Everett Reimer, "Modern Personnel Management and the Federal Government Service," The Federal Government Service, (New York: The American Assembly, 1954), p. 157.

⁴Senate Subcommittee on Federal Manpower Policies, Supervisory Selection in the Federal Government, (Washington: Government Printing Office, 1952), p. 7.

⁵Reimer, loc. cit.

the service.

In 1950, Civil Service Commissioner James M. Mitchell pointed his finger at what he called "the curse of excessive specialization."⁶ The historical emphasis on rigid job classification and technical speciality has extended unduly into the management area. Men starting out as technicians may develop into excellent executives, but technical proficiency in itself does not constitute fitness for administrative responsibility.

Furthermore, the fact that the top positions in the Federal Government are reserved for political appointees rather than career civil servants, while necessary to assure a measure of control to the party in power, has had adverse side effects. In addition to the desirable flexibility of policy that it is designed to achieve, this arrangement introduces unsought for discontinuities and tends to draw attention from the need for executive development within the career service.

Had not the Pendleton Act as passed by Congress in 1883 prescribed a civil service open in nature as opposed to a closed service with input only at the bottom, the

⁶James M. Mitchell, "Recent Progress in Federal Personnel Administration," Public Personnel Review, (October 1950) p. 181.

problem of executive development would necessarily have been faced much earlier. Nowadays with the continuing prospect of an unfavorable pay differential in the senior grades relative to industry, it is imperative that managers be systematically developed within the service.

The Second Hoover Commission stated:

Particular emphasis is needed on building a systematic governmentwide executive development program which will directly improve the quality of first line supervisors and junior executives, and which will also increase the supply of experienced and competent senior civil servants for top management posts.⁷

The approach to staffing in the Civil Service has, from the earliest days of the service, been oriented more toward the positions to be filled than toward the careers of the individuals. This is seen in the traditional concept that rank resides in the job, not in the person, which distinguishes the American Civil Service from the military services and the Civil Services of France and England.⁸ Instead of the question, "Whom can we get to fill this vacancy?" we should be asking, "Which are the outstanding employees and what can we do to develop them so that their potential can be most fully utilized in meeting the needs of the service?"

⁷Commission on Organization of the Executive Branch of the Government, Personnel and Civil Services, (Washington: Government Printing Office, 1955), p. 50.

⁸Paul Van Riper, History of the United States Civil Service, (White Plains: Row, Peterson, 1958), p. 299.

METHODS TO IMPROVE

It is seen, then, that the causes of management development problems in the Civil Service derive from its historical evolution. Although many of these causes and, to some extent, their effects lie beyond the power of individual agencies to rectify, present legal and administrative regulations allow sufficient latitude to the agencies to permit substantial improvement through their own efforts.

Any measures would contribute much if they did no more than to bring about a full awareness, throughout the organization, of the distinction between technical and administrative skills. It must be recognized that whereas technical competence in an organization is indispensable, it can be completely nullified through inadequate management. Improvement would result from a consistent application of the policy that a vital portion of every supervisor's duty must comprise the identification and nurturing of both types of skill in his subordinates.

Management training programs continue to be employed in government with varying degrees of success. Because conscientious supervisors often begrudge the time lost from regular duties on such programs, careful preparatory work is required to assure widespread appreciation of the

purpose and importance of the programs. They must be regarded as investments for the future.

Periodic rotation of executive personnel among different jobs and different divisions can provide greater opportunities for the acquisition of necessary management experience and practice. Here again, the short-run loss in efficiency can pay heavy long-run dividends in improved management and breadth of outlook.

Promotion selection methods can be improved within the scope of present civil service regulations. Changes in the actual systems can be introduced subject to Civil Service Commission approval. Some improvements can be effected through clear and emphatic agency-level policy guidance in handling present systems. For instance, emphasis can be shifted from the concept of promotion as reward for past services to one of promise of future services.

One of the most difficult, controversial and yet important elements of a management development program is the performance evaluation system. Some means of recognizing management potential and of recording the progress of supervisory personnel is indispensable in large organizations in order that prospective managers in different subunits be given something like equitable

treatment and that talent not be overlooked.

Discussing this subject in Personnel, two management consultants concluded "that the 'Achilles heel' of management development programs is to be found in the appraisal of performance."⁹ Optimum effectiveness of training, rotation and promotion programs can be attained only through the compilation and use of meaningful records of employee progress.

PERFORMANCE EVALUATION

Clearly, as Wallace Sayre said in his introduction to the American Assembly study of the government service, "No single formula will meet all the problems of the Federal Civil Service,"¹⁰ However, this paper is being written in the belief that the development and proper use of sound performance evaluation systems in government are indispensable to the much needed improvement in manager personnel. This area of study appears to be fruitful also for the reason that it offers much scope for improvement without changes in basic legislation.

It is true that no perfect system of evaluating personnel has ever been devised, but systems have been

⁹Walter R. Mahler and Guyot Frazier, "Appraisal of Executive Performance," Personnel, vol. 31, no. 5, (March 1955), p. 430.

¹⁰Wallace S. Sayre, Introduction to the Federal Government Service, (New York: The American Assembly, 1959), p.13.

developed which are workable, useful and reasonably fair. An imperfect tool is better than no tool at all. The objective judgement of an experienced supervisor, based on systematic observation of employees' abilities, attitudes and performance, is still the most reliable means of gauging present and potential value and readiness for greater responsibilities. Because human attitudes and behavior patterns are still less reliably measurable by other methods, the recommendation of a well-trained and competent manager regarding an employee should in most cases continue to weigh far more heavily than tests, interviews or any other criterion.

The main problem, of course, encompasses the entire realm of management staffing--of devising means of ensuring that all executive posts in the Civil Service be continually filled by managers of the highest possible calibre. The subproblem, to which this paper will seek at least a partial answer, is the question of what changes, if any, from present evaluation procedures are required in order to contribute most to the solution of the main problem.

Comprehensive analyses of Civil Service problems and far-reaching recommendations for change have been made in recent years by the American Assembly, both Hoover Commissions and before various Congressional committees. In

some instances for quite valid reasons, these suggestions have failed for the most part to obtain the necessary concurrence of Congress. The present study, therefore, will differ from previous ones, not only in its much narrower scope, but also in the fact that it will seek answers primarily within the limits of the presently existing legal and administrative framework.

In an effort to discover what, if any changes from existing practices may be desirable, the evolution of present systems and their effects in practice will first be described. These will then be examined in the light of recent research and industrial experience in order to develop possible alternative methods. Finally, the alternatives will be weighed and any resulting suggestions for improvement discussed.

SUMMARY

Reduction in the scope of government continues to be an appealing subject, particularly among the party which is out of office at a given time, but there appears to be little prospect for diminished federal government functions in the foreseeable future. Even those who advocate small government in principle often at the same time favor programs which would inevitably lead toward an expanded government role at the federal level. Any real reduction

in expenditures without abandonment of important national objectives can be achieved only through skillful and clear-sighted management.

Federal Civil Service personnel, despite the recently proclaimed cuts, will remain above the two and one-half million mark and thus will continue to exceed the manpower levels of any period of peacetime in its history. At the same time, organizational complexity and technological advancement mean not only that greater numbers of managers are required but that they must be more highly trained.

Many Federal agencies have been slow in recognizing and meeting this challenge. There has been a tendency to rely on one-shot formal training injections with insufficient attention to long range programs. Real improvement is a slow process involving, in addition to training, clear, persistent policies with top level support in the fields of recruiting, career planning, job rotation and promotion selection. The success of all these measures for executive development is dependent upon the existence and proper use of sound performance evaluation systems.

As a fundamental prerequisite to the development of managers, the area of merit rating in the Civil Service has been selected for study. Performance evaluation is a difficult and controversial subject, but a survey of the

rating practices of many agencies indicates that substantial improvements may be possible even within the state of this imperfect art. Fortunately, too, the present laws and regulations pertaining to the subject are sufficiently broad to permit more satisfactory solutions to many of the problems within the existing legal constraints.

CHAPTER II

PRESENT EVALUATION PROCEDURES

Requirements for speed and flexibility in meeting the expanding personnel needs of the federal government during World War II of necessity brought about a decentralization of personnel administration from the Civil Service Commission to the departments, bureaus and agencies. This trend has continued since the war, with many of the functions formerly performed by the commission now carried out in the personnel sections of the agencies in accordance with the policies and subject to the approval of the commission.¹

Until 1950, however, performance evaluation continued to be accomplished throughout the service in accordance with a single uniform system and using a single trait rating form prescribed by the commission. The system required the assignment of one of five adjectival ratings: Excellent, Very good, Good, Fair or Unsatisfactory. For promotion or an in-grade raise it was necessary that an employee's last rating be at least Good, and a rating of Fair or less normally called for dismissal proceedings.²

¹Paul P. Van Riper, History of the United States Civil Service, (Evanston: Row, Peterson, 1958), pp. 415-420.

²Felix A. Nigro, Public Personnel Administration, (New York: Henry Holt and Company, 1959), p. 307.

This system was severely criticized by the first Hoover Commission largely on the grounds that it failed to provide any means of improving employee work performance or of strengthening supervisor-employee relationships. The recommendations of the commission relating to employee appraisal called for:

1. Elimination of the adjectival ratings but continued annual appraisal of all employees in terms of work performance, improvement progress and growth potential.
2. Supervisor-employee discussion of ratings in which specific programs for individual development would be agreed upon and used to gauge progress.
3. Greater emphasis on training of supervisors in evaluation procedures.
4. Divorce of merit rating from formal personnel actions such as removal or award of in-grade raises. These latter would require separate certification by the supervisor in each instance.³

The task force also advocated that agencies should be permitted to participate in developing their own evaluation programs. Such participation, it was thought, would result not only in systems better adapted to the peculiar problems of each agency but in wider acceptance and enthusiastic support of the resulting systems. It was also pointed out

³ Commission on Organization of the Executive Branch of the Government, Personnel Management, (Washington: Government Printing Office, 1949), p. 40.

that the separation of ratings from personnel actions would obviate the cumbersome appeals apparatus.

These recommendations reflected the current thinking of leading personnel administrators of the time and were based on the up-to-date practices of industry. It was recognized that the total worth of a human being can not be summed up in a single word and that any attempts at such summations are apt to be neither useful nor meaningful to the individual or the organization.

A shift in emphasis was sought from lists of abstract and semantically indeterminate traits to more concrete elements of work behavior. Attention was to be directed away from the concept of the person as a static bundle of characteristics toward the concept that his progress toward agreed upon goals can be used as a more objective measure of his capabilities and potential. Last, the recommendations recognized that the achievement of goals satisfies a basic human urge and that knowledge of one's progress can be a spur to further achievement.

PERFORMANCE RATING ACT OF 1950

The Hoover Commission's study provided the impetus which resulted, among other measures, in the Performance Rating Act of 1950, the legal basis for the present evaluation practices in the Federal service. In the

enactment of this legislation, Congress, while seeming to concur with the objectives of appraisal expressed by the Hoover Commission, gave only partial endorsement to the means suggested for achieving them.

The Act requires that each department establish and use one or more rating plans, that the plans be as simple as possible and that each be approved by the Civil Service Commission as conforming with the act. It further stipulates that each plan provide:

1. that proper performance requirements be made known to all officers and employees;
2. that performance be fairly appraised in relation to such requirements;
3. for the use of appraisals to improve the effectiveness of employee performance;
4. for strengthening employee relationships; and
5. that each officer and employee be kept currently advised of his performance and promptly notified of his performance rating.⁴

The Act thus abolished the uniform rating system and permitted agencies to develop their own systems subject to approval of the Civil Service Commission. However, the close relationship between ratings and personnel actions was retained as was the concept of adjectival ratings.

⁴Public Law 873, chapter 1123, 81st Congress, 2nd Session. The text of the act also may be found in the Federal Personnel Manual, Chapter 21.

A minimum of three levels is required, Outstanding, Satisfactory and Unsatisfactory.

The provisions of the Act, though otherwise quite general, are very specific in circumscribing the use of these ratings. As unsatisfactory rating can only be assigned after a 90-day warning period and must occasion the removal of its recipient from his position. In order to limit the use of the outstanding rating, the Act stipulates that it may be awarded only if "all aspects of performance not only exceed normal requirements but are outstanding and deserve special commendation." All the legal implications formerly attached to "Good" of the former system were transferred to the present "Satisfactory" rating.

Instructions for establishing and administering performance rating plans in accordance with the new act were duly issued by the Civil Service Commission, and by mid-1951 over sixty different agency plans had been approved.⁵ These instructions, which comprise Chapter P-4 of the Federal Personnel Manual, adhere quite closely to the concisely worded provisions of the Act but offer additional guidance and amplification.

⁵Van Riper, op. cit. p. 435.

Whereas the Performance Rating Act calls for a minimum of three rating levels, the Commission prescribes a maximum of four, the optional additional level to denote performance between Satisfactory and Outstanding. The already rigorous constraints of the original Act over the award of Outstanding ratings are buttressed by the Commission's imposition of further requirements for elaborate documentation and review in each instance of the rating's use.

Few other restrictions beyond those of the Act are imposed by the Commission. Ratings are required at least annually for Grades GS-10 and below and not less frequently than every eighteen months for others. Agencies may not require that ratings conform to any predetermined distribution or use any method or process kept secret from supervisors or employees.

Except for the above, the details of the rating process are left very much to the discretion of the agencies. An agency may have as many different rating plans as it finds necessary in view of the nature of its component organizations and types of employee. Rating forms may vary considerably depending on the evaluation plan of the using agency. The use of forms is not even required as long as an official record of the performance ratings of each

employee is maintained, each is kept currently advised of his performance and notified of his ratings.

Among its guides for effective performance evaluation programs, the Commission suggests supervisor-employee discussions as being a device that has proved of value. Curiously enough, despite the Act's emphasis on the use of the rating process for work improvement and strengthening supervisor-employee relationships, these discussions are not specifically required to be part of the rating system.

The Commission also offers detailed suggestions for use in training supervisors in the most effective use of the rating program. Here again the mechanics involved and the degree of emphasis to be placed on training are left to the agencies, but each agency is required to describe its intended training methods at the time of submission of its evaluation program to the Commission for approval.

There is some justice in the common view of the Performance Rating Act of 1950 as a rather clumsy compromise between the "human relations" use of evaluation and its use for more mundane administrative purposes. The most obvious differences between the Act as finally implemented and the Hoover Commission recommendations are the retained summary ratings and automatic impact of ratings

on formal personnel actions. A more subtle difference may be seen in the Act's exclusive consideration of past performance of employees and complete omission of any mention of growth potential. Both in the wording of the Act itself and in the subsequent Civil Service Commission regulations, performance rating means past performance.

This is made very clear in the interesting final paragraph of the Federal Personnel Manual's chapter on Performance Evaluation, which describes the relationship between performance evaluation and other evaluation devices. These other devices are designed to serve personnel needs not strictly related to current performance such as:

- Promotability.
- Determining training needs and potential.
- Selection for executive development programs.
- Termination of the probationary period.

The Commission implies that, although the most recent performance rating should usually be one of the factors considered in making these decisions, other devices should also be used. Agencies are then admonished that a clear distinction should be drawn between performance ratings and the other devices and that, when used, the latter should be clearly identified as "not being part of an official performance rating plan approved by the Civil Service Commission." Thus, in a single sentence, the

Commission appears to absolve itself of further responsibility for or interest in what many regard as one of the most important functions of the evaluative process.

Only by the narrowest interpretation does the paragraph summarized above escape direct contradiction to the Performance Rating Act, Section 4 of which states:

No officer or employee of any department shall be given a performance rating, regardless of the name given to such rating, and no such rating shall be used as a basis for any action, except under a performance rating plan approved by the Civil Service Commission as conforming with the requirements of this Act.

In addition to performance ratings in the narrower sense, then, the Commission suggests that other evaluation devices are properly used by agencies in making certain types of personnel decisions. In practice, few agencies have ever employed such devices in any systematic manner. Either the attempt is made to include in performance ratings information useful in making these decisions or the decisions are made without any reference to systematically compiled data.

The numerous criticisms which have been directed against the systems resulting from the Act of 1950 can be classified in two major categories summarized as follows:

1. The use of ratings in promoting morale and efficiency has been hopelessly vitiated by the retention of adjectival ratings and their direct effect on personnel actions.
2. The use of ratings in determining wise and consistent personnel actions has been sacrificed to the false belief that evaluation processes, which have no other significance, can contribute materially to morale and efficiency.

The only agreement here is that the new programs have not been very effective for any purpose. Let us examine these two opposite viewpoints separately.

EVALUATION FOR MORALE AND EFFICIENCY

Despite the mandatory use and legal implications of the summary ratings, the primary intent of existing regulations governing rating systems appears to be the achievement of improved work performance and strengthened supervisor-employee relationships. Success in accomplishing these objectives is very difficult to assess.

In those work situations where it has been possible and the effort has been made to prepare task statements and performance standards for each job, some increases in efficiency have been reported. These measures, however, are time-consuming, and few agencies have applied them consistently. It is also questionable whether the increases attained have always justified the efforts expended. It should be pointed out, though, that lack of demonstrable

success in achieving greater efficiency can not in any large degree fairly be attributed to the summary ratings.

Any estimate of the extent to which supervisor-employee relationships have or have not been strengthened is even more of a guess. It is certain that these relationships have been very little affected in those agencies which have fulfilled only the minimum requirements of the regulation by periodic posting of machine listings of summary ratings. On the other hand, supervisor-employee post-rating discussions, when used, have resulted in instances of increased understanding, mutual respect, and successful efforts to acquire new skills and correct weaknesses. Granted, these instances have not been counted, and some of them would have occurred in any case. Considering the neglect of training in counseling, the task force of the Hoover Commission could scarcely have asked for more.

The point of all this is that except for a few obvious cases, it cannot be conclusively proven that the Performance Rating Act has failed in meeting its objectives of improved morale and efficiency. Even in the obvious instances it can usually be shown that the failings have resulted less from the shortcomings of the Act itself than from agencies' incomplete implementation of its provisions.

Nevertheless, the much-criticized adjectival ratings and the restrictions on their use have led to some curious results in practice.

Traditional rating systems are often criticized for the damage to morale resulting from "invidious comparisons" among employees receiving different ratings.⁶ One thing that can be said for the present system is that it has been quite successful in eliminating these comparisons--practically everyone now receives the same rating. The fact that ninety-nine per cent of all evaluations are now Satisfactory, however, minimizes their usefulness in determining the relative merits of civil servants and contributes to the complacency that ratings are designed to overcome.

To those who conceive the withholding or granting of within-grade pay raises as a legitimate and useful management tool for motivating toward good performance, the present legal requirements appear to constitute an undesirable limitation on the supervisor's prerogatives. Because a raise may be withheld only in the case of "unsatisfactory" performance, which also necessitates removal of the employee from his position, the step can no longer be taken in the case of those who are capable of performing well but for some reason fall slightly short

⁶For example see O. Glen Stahl, *Public Personnel Administration*, (New York: Harper and Row, 1962), p. 263.

of doing so. Under the present system, in-grade pay raises are, in effect, automatic. There is a growing belief among personnel administrators that automatic in-grade raises, if not preferable, are at least inevitable. It is universally agreed (except in Congress) that the administration of these raises ought not to be directly connected with performance evaluation.⁷

The restrictions upon and automatic consequences of the use of the Outstanding and Unsatisfactory levels unquestionably impose a limitation of the supervisor's ability to use ratings for motivating his subordinates. At the same time, however, under most systems now in use, he enjoys considerable latitude in recognizing merit and identifying needed improvement within the Satisfactory category. Although a small percentage of abuses may fairly be ascribed to the present status of summary ratings, it is doubtful that their abolition would, in itself, materially increase the effectiveness of performance evaluation for counseling.

Disappointing results in the use of present systems may actually be more attributable to the fact that the relationship between evaluations and subsequent personnel actions is not direct enough rather than too direct. When both the rater and the rated are fully aware that ninety-nine per

⁷For an interesting discussion on this, see "Personnel Opinions," Public Personnel Review, vol. 23, no. 3, (July 1962), pp. 201-205.

cent of the ratings are weighted little or not at all in personnel decisions, it is not surprising that performance appraisal is perfunctory and not taken very seriously by either party.

EVALUATION FOR MANAGEMENT INFORMATION

Although the success of performance rating in promoting morale and efficiency in the Civil Service is debatable, there exists no doubt whatsoever about its failure to provide any reliable indication of the relative abilities of employees. Few of the systems now in use attempt to achieve the objectivity and uniformity necessary for inter-unit comparison of employee qualifications, and few agencies make a consistent practice of referring to performance records in personnel actions other than to check the summary ratings.

Even granting that the production of such information is a subsidiary goal of present systems, its rare attainment is disappointing. Here again a sort of vicious circle is operating: because ratings are not consistently dependable, they are not used--because everyone involved realizes they are not used, insufficient care is taken in their preparation.

The adverse effects of the artificial restrictions over the use of summary ratings may be seen even more

clearly in this area than in the area of efficiency and morale. The fact that ninety-nine per cent of all ratings are satisfactory, of course, detracts from their usefulness to management and accounts for their being little used. However, even more serious is their effect in distorting the ratings themselves.

Consider, for example, the case of the research engineer who has achieved the technological breakthrough of the century but whose job description calls for breaking in new workers and briefing visitors. He must either be relegated to the ranks of the satisfactory or be certified by his supervisor as being an outstanding junior engineer-indoctrinator and visitor-briefer as well as an outstanding engineer. This sort of situation can and does lead to distorted appraisals. It is conceivable that this restriction might in some cases even exert an improper influence over the preparation of job requirements.

Many a truly outstanding employee is consigned to the satisfactory category because his performance of minor duties may be only a little above average. The final result often is less a function of the employee's qualifications than of his supervisor's scruples or perseverance in documenting the justification for the rating. Some agencies have further reduced the chances of objective

rating by pursuing the policy that the outstanding rating must carry with it a cash award.

By developing rating systems which permit discernment of the relative merits of "Satisfactory" employees, some agencies have succeeded in providing information used in promotion, training and assignment of personnel. However, because of the statutory force of the summary rating, finer gradations of performance and potential can not be considered in decisions relating to in-grade raises and reductions in force. The fact that nearly all employees receive the same summary rating is particularly harmful in the latter case. It means that job performance and future promise become less important to these decisions than seniority and veteran status.⁸ The heartbreaking result is that frequently the best prospects must be released in a personnel cutback while much of the dead wood is retained.

The failure of most agencies to develop any systematic approach to promotions, including the regular agency-wide review of the performance records of all eligibles, has resulted in inequities both to individuals and agencies. The present system, or rather lack of system, in addition to attaching excessive value to seniority, affords unfair advantage to those close to the centers of power. Staff personnel have progressed more rapidly than their line

⁸Van Riper, loc. cit.

contemporaries, due as much to their closer association with upper management as through their opportunities to gain the broader picture. Similarly head office personnel have tended to fare better in the promotion race than field personnel.⁹

Before concluding that present evaluation programs fail entirely in providing the data required for sound and equitable personnel decisions, it is necessary to examine the "other devices" mentioned so casually by the Civil Service Commission. What are these devices and how widely are they used?

A few agencies make good use of testing and training course results in making promotion decisions--mainly at the first line supervisory level.¹⁰ The use of testing, however, for the most part is confined to the filling of vacancies from outside the service.

The initiative in promotion actions often rests with the immediate supervisory and is exercised by his submission of a written (sometimes oral) recommendation. This

⁹Commission on Organization of the Executive Branch of the Government, Task Force Report on Personnel and Civil Service, (Washington: Government Printing Office, 1955), pp. 238-241.

¹⁰Senate Subcommittee on Federal Manpower Policies, 82nd Congress, Supervisory Selection in the Federal Government, (Washington: Government Printing Office, 1952), pp. 7-9.

recommendation rarely comprises a comprehensive analysis of strengths and weaknesses; it is an evaluation with a purpose and can be expected to present information selected with the purpose in mind. Supervisors often vary in their standards applied in making promotion recommendations, but frequently knowledge of a supervisor's standards is used in deciding on his recommendations. Nothing, however, can help the subordinates of the overly meticulous supervisor; recommendations cannot be acted on unless they are received.

Most commonly, promotion decisions are made by boards composed of supervisory personnel representing organizational segments varying in size. Where the segment is relatively small, the decisions are based on personal acquaintance with the candidates as well as the recommendations of their supervisors. Where a board represents a larger segment, its members are less well acquainted with the candidates, and the relative eloquence and influence of the supervisors become larger factors. The results of these deliberations are normally priority listings from which vacancies are filled as they occur.

In those cases where the promotion boards are drawn from and act for reasonably small segments, the operation has the advantage of being based on group judgment.

However, such fine segmentation results in differential promotion rates in different segments of the same agency. On the other hand, the functioning of boards based on large organizational segments frequently devolves into unwholesome competition among unit heads to obtain larger numbers of promotions for their subordinates.

This process bears a strong resemblance to election to club membership. A systematic consideration of relevant qualifications is rarely possible, board members not being able to spare sufficient time from regular duties and systematically compiled data normally being unavailable in any case. In some agencies, board actions tend to be geared more closely to time-in-grade completions than to the personnel needs of the agency, with the result that employees remain on the waiting lists for long periods.

The greatest weakness in this system is its tendency to promote employees within one organizational segment while better qualified candidates might be available in a neighboring division. An unsatisfactory effort to overcome this weakness has been attempted in some agencies by consolidating promotion lists, the actual interleaving of names being performed in accordance with an arbitrary formula in which seniority may be weighted heavily. Few

agencies of the regular Civil Service have developed agency-wide selection mechanisms based on the systematic review of either performance ratings or other devices.

SECOND HOOVER COMMISSION

The most sweeping condemnation of the effects of the Performance Rating Act is found in the report of the Second Hoover Commission. Although their study was made in 1954, performance rating in the Civil Service has not changed significantly since that time. The commission's final report stated baldly that "performance ratings are not in fact today a means of measuring the relative merits of employees." It further stated that the system "falls far short of the objective of creating a friendly and helpful human relationship between supervisor and employee." It described the process as burdensome, perfunctory and requiring an expenditure of time and effort disproportionate to the benefits obtained.

These conclusions are uncontroversial enough, but the solution suggested by the Commission is surprising. Reasoning that good supervisors will induce efficiency and harmonious work relations without a formal encouragement-on-schedule program, they recommended that the existing performance rating system be replaced by one designed entirely to serve management needs in taking formal

personnel actions. Specifically the report recommended:

- (a) The present performance rating system be abolished;
- (b) A new system be established under which the supervisor reports at least once a year only on those:
 - (1) Employees with potential capacity for further development and higher responsibilities;
 - (2) Employees deserving meritorious awards;
 - (3) Employees miscast in their present assignments and who should be reassigned or retrained for other work;
 - (4) Employees undeserving of periodic pay increase because of unsatisfactory service (employees otherwise would receive their raises in due course);
 - (5) Employees requiring dismissal.
- (c) In each case the supervisor should submit in writing to his superiors the reasons for his judgment and specific suggestions as to action desired.
- (d) The right to appeal adverse decisions under the new system should be limited to one appeal to higher authority in the agency.¹¹

Although these recommendations have never been adopted, during the six years that elapsed between the reports of the two Hoover Commissions the pendulum had swung from a human relations view of performance evaluation completely

¹¹Commission on Organization of the Executive Branch of the Government, Personnel and Civil Service, (Washington: Government Printing Office, 1955), pp. 63-65.

over to a call for exclusive management use of ratings.

SUMMARY

The large degree of latitude granted by the Civil Service Commission to agencies in implementing the Performance Rating Act of 1950 has resulted in a wide variation in the evaluation systems employed in different agencies. Practices range from the periodic marking of simple summary ratings on machine punch cards at one extreme, to elaborate scale-rating against individually prepared lists of task elements complete with post-rating discussions at the other.

With few exceptions, these systems have failed to realize their full potential either for enhancing performance and work relationships or providing information for management use. Slight gains in achieving the former have been attained only at a high cost in terms of the latter.

Legal and administrative stipulations as to the assignment and impact of the summary ratings have had undesirable consequences, some of which are surmountable within present constraints and some of which are not. These may be summarized as follows:

Surmountable:

1. The large proportion of satisfactory ratings

minimizes their utility in personnel decisions;

2. The large proportion of satisfactory ratings leads to perfunctory evaluation and complacency;
3. Supervisors' ability to motivate is restricted;
4. Ratings are sometimes influenced by irrelevant factors;
5. Seniority may be excessively weighted in promotions.

Insurmountable without Legislative Change:

1. In-grade raises are essentially automatic;
2. Performance is a negligible factor in reductions in force.

Many of the shortcomings in the existing evaluation systems are attributable to a lack of clearly defined, well understood and consistently executed promotion and performance rating policies in the different agencies. Also notable has been the inadequate training of the raters.

CHAPTER III

DISCUSSION OF ALTERNATIVES

If an observant reserve officer spends a two-week training tour in a branch of a Civil Service-manned defense agency, he begins to develop a fairly accurate idea of which employees are competent and which are not. He can identify effective supervisors and productive workers, and he can distinguish the mere time servers even in that short a period. Yet, if he were to assemble the performance ratings of all the personnel of the branch, he would be hard put to discover any correlation whatsoever with his own observations.

There are several reasons for this. First is the fact that the avowed purpose to which Civil Service ratings are put are often not consistent with objective reporting. Second is the natural reluctance of supervisors, even when objectivity is sought, to be completely objective in a record which the employee will see. This reluctance stems in part from an irresponsible desire to avoid awkwardness but also from the psychologically sound realization that coldly precise employee analysis under these circumstances rarely produces improved performance and often has the opposite effect.¹

¹Kenneth E. Richards, "Some New Insights into Performance Appraisal", Personnel vol. 37, no. 1, (July, 1960), p.28.

A further source of unevenness in the present practice is an inadequate understanding on the part of supervisors of the techniques and significance of the evaluation process. Inconsistencies creep in because raters have not been instructed as to the objectives of the system used in his organization. In order to discuss rating systems intelligently it is first necessary to examine the purpose they are meant to serve.

PURPOSES OF EVALUATION

From a review of literature on the subject, it would be possible to compile a list of fifteen or twenty different purposes for which rating systems are used, but these can be consolidated with little loss into two primary areas:

1. To bring about improved work and work relationships through feedback to both raters and rated.
2. To provide information for management use in personnel actions such as promotion, transfer, removal and training.

The first of these objectives is normally accomplished by discussions at or near the time of evaluations, in the course of which closer agreement on work standards may be reached, suggestions for improved methods may be forthcoming and prospects for the employee's future development may be examined. Properly handled, these discussions can provide an opportunity for both supervisor and employee to gain a better understanding of the other's problems and viewpoints.

Potential trouble areas can be eliminated before they become serious. Recognition is given for good work, and the employees realize that seniors are really interested in their welfare and in them as people.

The second purpose, on the other hand, requires a detached objectivity on the part of the supervisor. Ratings, if used for this purpose alone, would properly provide a basis for comparison of present and potential capabilities of different employees. Estimates of both weaknesses and strengths would be frankly stated. This calls for a thorough, consistent and systematic approach applied uniformly in the evaluation of every subordinate. It requires standardization insofar as possible, in the methods and criteria applied by all reporting seniors.

It is seen, then, that these two broad objectives are not perfectly compatible in a single system. A system which serves one well does badly by the other. Maximum achievement of the first objective requires a supervisor to apply different treatment with different subordinates. Some workers benefit from an occasional jolt, whereas most respond best to encouragement. It would be brutal and fruitless to point out to a subordinate faults which lie beyond his power to correct and which are tolerable in his present job. Best results are gained, therefore, by careful

selection of points to be discussed with an employee without attempting to be systematically thorough. All of these factors, then, introduce inconsistencies which are undesirable from the viewpoint of the second objective. At the same time, the detached objectivity necessary for the second purpose would in many cases adversely affect both work performance and personal relations.²

Largely because of this incompatibility and based on the assumed use of only one evaluation system, the modern trend is to emphasize the paramount importance of work improvement and increased worker-supervisor rapport. This is reflected in recommendations of the First Hoover Commission and the current writings of virtually all authorities on personnel administration. They urge that performance evaluations ought not to be used as the basis of any formal personnel actions whatsoever.

The conclusion seems justified that the work improvement and strengthened relations benefits of rating programs are enhanced if the formal action use is discarded. Yet these authorities are for the most part silent on the subject of what alternative methods should be employed for

²A. S. Glickman, "Effects of Negatively Skewed Ratings on Motivations of the Rated," Personnel Psychology, vol. 8, no. 1 (Spring, 1955) p. 41.

the latter purpose. The view seems equally justified that the original purpose of performance evaluation is still a valid requirement. Some systematic method is needed to ensure that talent be recognized by the organization as a whole and put to its most effective use.

Even in a small unit, the selection of the most suitable man to fill a higher position requires careful consideration. In large organizations such as our federal agencies, employing in some cases thousands of people, the recognition of those best qualified for advancement presents an extraordinarily challenging problem. It is particularly difficult in the absence of any reliable means of comparative evaluation of candidates.

One of the easiest and a reasonably satisfactory method is to decentralize promotion actions to lower organizational levels. This, in effect, breaks a large entity down into smaller, more manageable pieces in which decisions can be made by those who are acquainted with the candidates. This is in accordance with current Civil Service Commission philosophy and corresponds with a recommendation of the second Hoover Commission.³ This policy is superior to the cumbersome, overly centralized

³Commission on Organization of the Executive Branch of the Government, Personnel and Civil Service, (Washington: Government Printing Office, 1955), p. 85.

personnel administration which preceded it, but it suffers from two serious defects. One is the tendency, when a vacancy occurs, to fill it with whomever is available, no matter how unpromising. The other shortcoming is the related result that mediocre employees in one segment may advance more rapidly than promising ones in another. The alternative is to broaden the base for selection over a wider organizational area, and from this arises the continued need for a reliable means of comparing the relative qualifications of candidates who are not well known to the selectors. The evaluation systems presently used in most federal agencies fail to provide such a means.

Should anything like a unified career Civil Service, such as recommended by the second Hoover Commission, ever be approached, the inadequacies of the present evaluation systems would become immediately apparent. Such a service, which envisages freer rotation of executives among agencies, presumably would benefit also from a service-wide promotion program, to which a uniform and truly meaningful evaluation system would be indispensable.

Until a unified service is achieved, and probably even afterwards, significant advantages may be gained through the use of different merit rating systems for different purposes. The personnel problems of the different

government departments and agencies vary widely in nature. Some elements contain a high proportion of repetitive tasks which easily lend themselves to quantitative and qualitative measurement. Others comprise office work less easily measurable. Some agencies employ large numbers of engineers and technicians, a situation which produces its own unique management problems. Evaluation systems could be tailored to the peculiar needs of each agency, and some agencies might beneficially employ several different systems. In every case, the purposes to be served by the systems must be agreed upon and kept clearly in view both by those who develop and those who use the systems.

EVALUATION FOR PROMOTION

One close observer of the federal Civil Service has written, "One of the objections to rating systems in the past has been focused on the effort to serve too many purposes with a single tool."⁴ The solution may be to acquire the necessary additional ones. It is tempting to suggest the use of parallel systems--one for use in feed-back to the employee and the other solely for management's information. This is not a new idea, and different variations of it have long been used in industry.

⁴O. Glen Stahl, Public Personnel Administration, (New York: Harper and Row, 1962), p. 261.

A General Electric employee relations manager has described the shortcomings of one such system used until 1954 in a division of that company. The system employed two trait-rating forms, one of which was used for discussion with employees. Ratings on the other form were used for salary administration and were not disclosed to the employees. Not surprisingly, difficulties arose in cases where employees, having been told their work was good, failed to get raises.⁵ Similar difficulties are likely to be encountered wherever salary administration is based on factors apart from the openly disclosed evaluation system. Regardless of the rating system employed, these difficulties can be avoided only through full understanding and the exercise of consistent good judgement on the part of supervisors.

Difficulties of the above type are encountered most frequently where, as in the Civil Service, seniority weighs too heavily in promotion decisions. Wherever promotions regularly occur soon after expiration of minimum time in grade, selection on real merit runs counter to employee expectations and occasions bitter feelings, particularly if the evaluation process is performed

⁵M. S. Kellogg, "Appraising the Performance of Management Personnel," Personnel, vol. 31, no. 5, (March, 1955), p. 443.

indifferently. On the other hand, in organizations where a consistent and well-understood policy of promoting only the outstanding is followed, good job performance does not in itself necessarily imply fitness for promotion. In case of question, it is then possible to point out that those promoted had better qualifications.

Another device commonly used in industry is to apply different evaluation methods to different categories of employees. The performance of first echelon workers, for instance, can often be appraised quite accurately in terms of quantity and quality of output. The effectiveness of supervisory personnel, on the other hand, is less easily measured and involves a completely different set of criteria. The single, very broad rating form presently used in many federal agencies is not well adapted to both purposes. Some agencies might benefit from the introduction of a separate rating form for each of these two categories. The dividing line might be based on GS level, depending on the composition of the work force in an agency and the nature of its work.

In activities employing significant numbers of engineers or other technical specialists, categorizing by grade level is not sufficient. For these people, adequate promotion opportunities must be provided which do not

require assumption of managerial responsibilities. This can best be done by treating executive and technical promotion streams separately, including the evaluation systems used. This does not imply that any restrictions should be placed on the entry of properly qualified technicians into the executive stream; it simply means that advancement must also be possible for those technically qualified but who are either not interested in or suited for administrative work.

Some authorities object to the use of performance ratings in promotion actions on the ground that good performance in one position does not imply qualification for a higher post.⁶ Certainly the management development problems in the federal service have been aggravated by the undiscriminating selection of supervisors purely on the basis of technical competence. However, when the question is one of increasing the responsibilities of those already occupying supervisory billets, previous performance is not only valid; it is the most reliable of all bases for action. The use of a separate form for appraisal of supervisory skills would help to focus the attention of the reporting senior on the special qualities contributing to good

⁶Thomas L. Whisler and Shirley F. Harper, Performance Appraisal, (New York: Holt, Rinehart and Winston, 1962), p. 434.

management performance.

The point is sometimes raised that it is difficult to assess an employee's fitness for promotion without knowledge of the specific job to which he might be advanced. This is true in general, but less so as applied to administrative posts. The problem might be overcome by the submission of special evaluations for all candidates whenever vacancies occurred. This would be a useful procedure after the field had been narrowed to a small number of candidates but would be unduly burdensome if applied to a wide field. More routine evaluation procedures would still be required as a basis for narrowing the field. A more feasible means of minimizing this problem might lie in the use of the separate forms for supervisor evaluation previously suggested. Many agencies lay out tentative routes of advancement, which can be of assistance to supervisors in preparing their evaluations. It may also be helpful to require seniors to suggest one or several jobs to which they think each subordinate is suited as a part of the regular periodic evaluation.

CONFIDENTIAL REPORTS

It was earlier suggested that a major trouble area in performance evaluation arises from the fact that it is difficult to accomplish the two broad objectives of merit

rating by a single system. It was shown that consequently the present trend in government and progressive industry is to emphasize the work improvement and supervisor-employee relations aspect of the appraisal process at the expense of information to be used for formal personnel actions. The opposite approach would be to confine the evaluation process to management's use, taking the view that good morale and employee development are properly part of the supervisor's year-round responsibility. It would then no longer be necessary to disclose the ratings to the rated.

One result would be an increased objectivity in reporting. This was demonstrated in a study at Lockheed in 1949 in which 485 supervisors raised their average grades from 60 to 84 on a scale of 100 as a result of imposition of the requirement that ratings be made known to employees.⁷ Presumably also, employees would be spared the distress of invidious comparisons, and less time would be lost in appeals proceedings.

Such a practice is followed in the British Civil Service, where subordinates never see their painstakingly prepared annual reports. The concept of confidential

⁷Lee Stockford and H. W. Bissell, "Factors Involved in Establishing a Merit Rating Scale," Personnel, vol. 26, no. 2, (September, 1949), pp. 94-116.

reporting also is approached in the American Navy's rating of its officers in that satisfactory reports are not normally shown to the officers reported on at the time of their submission.

Under the British system, promotions depend on the deliberations of an annually convened, three-man promotion board, before which each eligible candidate appears in person with his own supervisor present in the role of "prisoner's friend." Evaluation reports are also carefully considered, but an individual who "boards well" tends to be chosen over one whose reports are equally good but who fumbles his interview. On the other hand, a candidate whose supervisor has marked him "extremely well qualified" is reasonably assured of promotion (although he doesn't know it, of course) despite nervousness, indifferent grammar or otherwise unimpressive showing before the board.

The effectiveness of this system would be much reduced through excessive use of the EWQ rating. Moderation in its use is achieved, not by administrative barriers as with the outstanding rating in the American service but by the understanding and appreciation of the system by reporting seniors. The quality of reports is also enhanced by critical review at echelons above that of the original rater.

Confidential appraisals have undoubted merit when applied by supervisors who conceive the development of subordinates as an important duty to the parent organization. However, they do not automatically provide the opportunity for periodic employee self-appraisal nor, in themselves, contribute to good supervisor-employee understanding. Moreover, they are subject to abuse by unintelligent or biased supervisors, there, of course, being no appeal. Even a mandatory provision for informing the employee of unsatisfactory ratings does little to overcome this shortcoming, because, in practice, a mediocre rating can often blight a career as effectively as an unsatisfactory one.

Secret rating systems also are contrary to the principles of open disclosure normally looked for in the operations of our government. The Civil Service Commission has prohibited the use of any "rating method or process" which is kept secret from either supervisors or employees.⁸ Whether or not this prohibition would extend to the results of an openly described method or process is not clear, but it is highly unlikely that any system resembling that of the British service would meet with the approval of either the

⁸U. S. Civil Service Commission, Federal Personnel Manual, p. P-4-4.01.

Commission or Congress.

Although use of performance evaluations for feedback to the employee undoubtedly detracts from the accuracy of the information management should have as the basis of personnel actions, no other alternative appears practicable for application in the Civil Service under the present law. Effort must be devoted to the development of systems which serve both purposes as well as possible. Survey of industrial experience with appraisal systems indicates that this is possible, but success in its achievement lies less in the mechanics of the systems than in the attitudes of supervisors toward the job of performance ratings. It has worked best where they recognize the importance of searching for the most promising promotable material among their subordinates.

METHODS OF EVALUATION

In developing its merit rating program, an agency is constrained by the Performance Rating Act of 1950 to assign the single, overall adjectival rating of outstanding, good or unsatisfactory or, optionally, a fourth level between good and outstanding; however it is free to employ as many rating systems as necessary, and it is granted considerable latitude in developing rating forms. The Civil Service Commission has indicated its willingness to approve forms

which provide greater detail than afforded by the simple adjectival ratings required by law. The Federal Personnel Manual suggests either narrative statements or marking against lists of characteristics or job requirements, either predetermined or of the fill-in type.

Few departments have fully exploited the choices open to them. Even now the formal evaluation program of some agencies is limited to the publication of IBM listings of adjectival ratings, satisfying the minimum requirements of the law but providing neither meaningful feedback to employees nor information to management. Others make use of narrative statements, which are useful in discussions with employees but do not ensure systematic consideration of all aspects of performance and are not very helpful to higher management. The most significant conclusions about employee potential that can be gleaned from such statements often lie not in what they say, but what they fail to say.

Graphic rating scales are widely used both in and out of Government. Properly developed and used, they can be very effective for a fairly wide range of purposes. Objection to their use has been aimed at their being subject to "halo" effect and the tendency for marks to pile up at the high end of the scales.⁹

⁹Erwin K. Taylor and Roy Hastman, "Relation of Format and Administration to the Characteristics of Graphic Rating Scales", Personnel Psychology, vol. 9, no. 22, (summer, 1956), p. 187.

In an effort to overcome the latter shortcoming, forced distribution techniques have been developed and, to a limited extent, used in industry. These techniques have validity as applied to single groups, but, in common with paired-comparison and rank-ordering schemes, are poor devices where inter-group comparison is desired. Forced distribution systems, moreover, are prohibited by present regulation.¹⁰

In order to minimize halo effect and semantic variances, the army has experimented with a forced choice method of evaluation, for which claims of high validity have been made. In this system the rating form consists of a number of groups of four statements each. The rater is required to select from each tetrad the statement he considers most true of the subordinate and the one he regards as least true. This system gains its effectiveness from the fact that the rater is not supposed to know which of the statements are significant to the scoring. Studies have indicated that this method may offer some gains in objectivity over some of the more conventional systems. These slight gains, however, do not appear to outweigh the disadvantages of expensiveness, complexity, difficulty in detecting errors, low acceptance by raters and

¹⁰U. S. Civil Service Commission, loc. cit.

unsuitability for discussion with employees.¹¹ These systems also appear to fall under the previously mentioned Civil Service ban against methods kept secret from supervisors.

The Critical Incident rating technique used by General Motors has attracted favorable attention. This method strives to compensate for the increasingly recognized unreliable memory of supervisors and to minimize the judgment element in ratings through the systematic recording of all significant behavior incidents at the time of occurrence.¹² It is claimed that the results of this process will be purely factual, but it appears that some subjective element may be retained through variations in the interpretation of incidents. The process, if applied to rank and file, could be time consuming, but it serves to focus supervisory attention on the continuous nature of his appraisal responsibilities.

The day may not be far off when management will bring itself to rely at least in part, on peer ratings or ratings of supervisors by subordinates. The fairly frequent use of peer ratings for the validation of other rating techniques is based on the researchers' belief in their

¹¹Lee W. Cozan, "Forced Choice: Better than Other Rating Methods?" Personnel, vol. 36, no. 3, (May-June, 1955), p. 83.

¹²John C. Flanagan and Robert K. Burns, "The Employee Performance Record," Harvard Business Review, (September-October, 1957), pp. 95-102.

greater validity. Some of this validity stems from the fact that such ratings are group judgments, but it also recognizes the fact that it is usually easier to deceive one's supervisor than one's colleagues or subordinates.

Peer ratings have been dismissed by some as mere popularity contests.¹³ It is generally less recognized, but the same criticism can be directed almost equally well at the conventional systems. The Lockheed study previously mentioned revealed a significant negative correlation between validity and length of rater-worker acquaintance, even in the ratings assigned by trained supervisors.¹⁴

Peer or subordinate ratings presumably would be used as supplementary evaluation systems and would thus represent additional, fairly considerable expense. They obviously are subject to abuse, but perhaps not much more so than conventional systems. The fact that they provide averages would tend to lessen the effect of extreme markings by a minority.

It should be noted that these systems call for anonymity on the part of evaluator. Their use up to the present has normally been with the stipulation that they

¹³ Robert J. Wherry and Douglas H. Fryer, "Buddy Ratings: Popularity Contest or Leadership Criteria?", Personnel Psychology, vol. 2, no. 22, (Summer, 1949), p. 147.

¹⁴ Stockford and Bissell, loc. cit.

would not form the basis for any official action. As such, they have provided useful information for supervisor self-improvement and offer an intriguing area for continued experimentation.¹⁵ Their use for more official purposes may be questioned on the grounds that the raters can not be held responsible for the results.

In studying the more recently developed appraisal techniques, the enquirer finds himself on decreasingly solid ground. Many devices already in use are based purely on theory or intuition. Most of the experiments with new methods have depended on correlations with results of other evaluation techniques of equally questionable validity. The most reliable criterion for validation of supervisor evaluation techniques is the long-run productivity of subordinate groups. The costs in time and money incurred in collecting enough of this kind of information to yield significant results are usually prohibitive. Any other approach to validation ultimately depends on somebody's judgment, either individual or collective. The assumption on which most validation rests is that a consensus is more accurate than an individual judgment. This is probably a good assumption, but is not an unassailable fact.

¹⁵P. W. Maloney and J. R. Hinrichs, "A New Tool for Supervisory Self-development," Personnel, (July-August, 1959), pp. 45-53.

Probably only because it has been longest in use, the most completely and satisfactorily tested evaluation method is the graphic rating scale. Both its advantages and disadvantages are well understood. Better systems may already have been devised, but none has yet been proved to surpass it for simplicity, accuracy and versatility. Carefully prepared systems of this type, used by trained and conscientious raters, have demonstrated their capability of producing usable information, acceptable to both the raters and the rated.

EVALUATING THE EVALUATOR

Much time, money and effort have been spent by both government and industry in the search for better methods of evaluating past and predicting future work performance. The one consistently recurring conclusion seen in all these trials and experiments is that successful results depend more on the raters than on the mechanics of the system. One result of this shift in attention to the rater has been wider use of training in evaluation. Every organization must decide how much time and money it wishes to invest in these programs, and estimation of their payoff is very difficult. It may take years for the resulting benefits to be felt.

It has been demonstrated, however, that training produces more valid evaluations. Admittedly, "valid" in this sense means similarity in appraisals of the same subordinate by different raters, but this in itself is a worthy goal.¹⁶ It is achieved by developing among raters greater uniformity in understanding of rating forms, approach to rating problems and concepts of normal work performance.

The Civil Service Commission requires all agencies to provide training in the use of its rating system, but in most cases this has been very perfunctory, usually consisting of a periodic routing to supervisors of the appropriate instructions.¹⁷ This weakness is aggravated by the fairly common practice of treating evaluation as a transaction strictly between the rater and the personnel department with little interest shown by the rater's line superior. The burden and expense of formal training in rating can be reduced by the regular and critical review of supervisors' rating efforts by their own supervisors.

Performance evaluation becomes a meaningless ritual unless the supervisor knows:

¹⁶A. G. Bayroff, Helen R. Haggerty and E. A. Rundquist, "Validity of Ratings as Related to Rating Techniques and Conditions," Personnel Psychology, vol. 7, no. 1, (Spring, 1954), p. 112.

¹⁷Federal Personnel Manual, p. P-4-4.

1. That his ratings are used, not just filed away.
2. That management as a whole and his own superior in particular are serious about the rating program and interested in his ability as a rater.

Substantial improvement in the effectiveness of rating systems can be attained if every executive is given to understand that he is not only responsible for his own ratings, but also bears a responsibility for the ratings assigned by his subordinates. This implies that review of all ratings should be required at the echelon immediately above the rater and that all supervisors should be evaluated on their performance as raters. There may be even more to be gained from this policy than increasingly valid performance appraisals. Fiedler's work suggests a positive correlation between leadership success and ability to discern dissimilarities in subordinates.¹⁸

A further shortcoming of the present system, as it is used in many agencies, is the practice of rating each employee on "anniversary dates", usually on or about the anniversary of his achieving his present grade or appointment. Although the Civil Service Commission stipulates only minimum frequency of reports, leaving timing to the discretion of the agency, the above practice is fairly

¹⁸Fred E. Fiedler, "Assumed Similarity Measures as Predictors of Team Effectiveness," Journal of Abnormal and Social Psychology, vol. 49, (1954), pp. 381-388.

widespread. Its purpose is to distribute the supervisor's evaluation task uniformly over the year, but its effect is to minimize the possibility of comparing the performance of subordinates of the same grade as part of the evaluation process. A good compromise solution to this problem is to stagger the evaluation period of the different grades throughout the year. This method is being used successfully for the submission of naval officers fitness reports.

SUMMARY

The two primary purposes of performance evaluation, employee counseling and information for personnel actions, are not easily compatible in a single system. Techniques well suited to one purpose can only with difficulty be adapted to the other. Systems have been devised which, when used by trained raters and when subject to moderate controls, have accomplished both purposes reasonably well. However, these systems perform neither function quite as effectively as systems designed to serve one function or the other alone.

Programs which attempt to accomplish both purposes separately by the simultaneous use of two independent systems consume much time and effort and encounter the problem of inconsistency of reports under the two systems.

Organizations, therefore, sometimes tend to emphasize achievement of only one of the purposes in the design of their formal rating systems hoping to achieve the other by informal means. Examples are seen in the present emphasis on counseling in the American Civil Service and the opposite concept as represented by the confidential annual reports of the British service.

If only one of the common purposes of appraisal is to be served by the formal system, it appears that the counseling function should be excluded. A standardized treatment of counseling is often less successful than a varied approach, whereas standardization is the very essence of comparative evaluation. Some authorities view evaluation primarily as means of systematizing a judgmental process which occurs in organizations whether or not a formal system exists and which requires standardization to avoid unfairness both to the employees and the organization.¹⁹

Many organizations have increased the effectiveness of their appraisal programs by adopting appropriately different methods for use in rating different categories of employees.

¹⁹Wallace H. Best, "Some New Directions in Personnel Appraisal", Personnel, vol. 34, no. 2, (September, 1957), p. 46.

Of the several more recently developed rating techniques, the critical incident method and some forms of group rating appear to be suitable for use in the Civil Service. For reasons of economy, employment of the latter would normally be limited primarily to higher management personnel. Free-written ratings are particularly appropriate for use in counseling but are not, in themselves, sufficiently systematic for comparative use of ratings. The great versatility, simplicity and reasonable accuracy of well-prepared graphic rating scales commend them for wider use.

The effectiveness of rating programs can be greatly enhanced by top management support and promulgation of clearly stated and often reiterated policy. This policy can be strengthened by providing training in the use of evaluation programs and by requiring critical review of ratings by the rater's line superior.

CHAPTER IV

SUGGESTED IMPROVEMENTS

Despite its defects and limited success in practice, the Performance Rating Act of 1950 offers one great advantage. The considerable freedom left to agencies in developing their rating systems undoubtedly can be blamed for some of the indifferent results, but, at the same time, it permits correction of many of the deficiencies without further legislative changes. The alternative of imposing more rigid requirements on all agencies would produce evils of a different kind and is not recommended. At the same time, the widely varying nature of personnel problems encountered in different agencies makes it possible to recommend few corrective measures that would be universally applicable.

One broadly applicable recommendation, however, is that every agency should periodically assess the adequacy of its personnel program. Such assessments would often reveal undesirable discrepancies between policy and actual practices, and can also serve to bring about policy changes necessitated by changing conditions. Each agency should assure itself that the design and operation of both its performance evaluation and personnel development programs are consistent with each other and with the needs of the

of behavior which are indicative of qualification for higher duties.

The most widespread weakness of present systems is their failure to provide consistently reliable information for management use in determining training needs, career-planning and promotion. The virtual exclusion of this extremely important function from the formal systems encourages the use of informal processes with consequent undesirable variation in criteria and neglect of potential. Efforts should be directed toward the development of formal evaluation methods which would permit a more systematic accumulation of this kind of information about employees, and the resulting performance records should be regularly referred to in making personnel decisions. This closer coordination between evaluation and selection for training, transfer and promotion would restore a sense of reality and purpose often lacking in the present systems.

The counseling function of performance rating should be retained and strengthened. The Second Hoover Commission, in its examination of existing practices, appears to have made a reasonably accurate diagnosis, but the remedy it suggests is more severe than is warranted by the disease and underrates the value of the counseling aspects of performance evaluation.

agency. This also permits focusing the attention of all raters on the purposes of evaluations, which prevents the process from devolving into a routine, bothersome and meaningless administrative task. Some of the needed changes brought out by these assessments would require prior approval by the Civil Service Commission. On the other hand, much of the needed improvement will be found to be attainable merely through more effective administration of present programs.

Many Civil Service-manned Federal agencies still are devoting inadequate attention to the recognition and development of executive potential within their organizations. Efforts to rectify this deficiency should include increased emphasis on the construction and utilization of performance evaluation systems specifically designed to aid in the identification and cultivation of administrative skills.

The present haphazard and perfunctory rating practices in many agencies often fail to provide either the basis for employee self-improvement or information useful to management. Frequently the attempt is made to apply identical rating procedures to rank and file, technical specialists and executives. Consideration should be given to the development of separate rating systems for each category, designed to direct the rater's attention systematically to those aspects

The requirement for a periodic stocktaking of the performance of each subordinate, when approached in the right spirit, can and does result in work improvements, the possibility of which might otherwise be overlooked even by the model supervisor. However, this is a tool requiring great skill on the part of the user, which explains both its frequent lack of success and supervisors' lack of enthusiasm for the process. It does not follow that skill in the use of evaluation for counseling is unattainable nor even that it is so rarely attainable as to invalidate the entire concept.

Benefits from employee counseling are rarely quantifiable and it is probable that the counseling use of performance appraisal has been more successful than generally realized. It is certain, however, that the successes have not approached the reasonably attainable, and the Hoover Commission's assertion that the evaluation process now takes more out than it puts back into the organizations may be a fair one. Such a conclusion is hardly surprising in view of the widespread misunderstanding of the purpose of the process and the insufficient appreciation of the skill required for its effective use. The disappointing results should be identified with these two factors, not with the process itself.

In order to attain full benefits from post-evaluation counseling, most agencies will find it necessary to make even a further investment in the form of rater-training programs. It is extremely difficult to weigh the relative costs and gains associated with evaluation counseling, but some reassurance is gained from the fact that many profitable corporations regard these programs as economically sound, at least as applied in management development. It also seems likely that the present deemphasis on these programs at the lower levels of industrial type installations of the Civil Service is the correct economic decision.

Having decided that both major functions of performance appraisal fulfill important requirements, the most difficult decision must then be faced--whether or not the attempt should be made to fulfill both requirements in a single process. It has been established that the attempt to do so involves some sacrifice in the adequacy of the results for either purpose. Nevertheless, for reasons of economy and the lack of any very satisfactory alternative, it is believed that combining the two functions would prove to be the most satisfactory compromise in most agencies. Reasonably good accomplishment of both functions in a single system, which has been shown to be attainable in industry, is preferable to neglecting one function or the other.

In order to provide useful information for management, consideration should be given to the development of more detailed rating forms. The intent here is to obtain systematic assessments relative to predetermined standards applied to every element of performance regarded as essential for success in the type of work being appraised. Whereas free-written appraisals are valuable, particularly in counseling, systems can be improved by the addition of graphic scales carefully prepared for each major job category as appropriate. The initial preparation of these forms, if properly done, is a fairly sizable project; but, once prepared, they facilitate the actual rating process. Use of such forms would not only provide more reliable information on individual strengths and weaknesses, but would also permit a finer sorting of overall markings among the ninety-nine per cent who are now graded "Satisfactory".

As a further step toward standardization, it is suggested that the practice of evaluating on anniversary dates be replaced by schedules which would require appraisal of all workers of the same grade and work category at the same time. The rating work load could still be distributed over the year by staggering the due dates for different grades and categories.

Even the most brilliantly devised performance rating program will fall short of its objectives if its users fail to understand it. In order to realize the full benefits of the program, it is necessary that raters, reviewers and users of the information all have as nearly as possible the same understanding both of the standards involved and the purposes of the system. This can only be achieved through carefully planned and repetitive training. Training in the use of existing evaluation systems is commonly slighted for reasons of economy. Yet, the past disappointing results from these programs usually can be traced to inadequate understanding of the rating process. If evaluation programs are worth having at all, they are worth the small additional investment required to make them effective.

Once a satisfactory understanding of the rating systems has been reached among present supervisors, the training burden can be eased and uniform standards maintained by insistence on the regular, critical review of each supervisor's ratings by his own superior. By adherence to this policy and by requiring that each supervisor himself be rated on his performance as an evaluator, the presently lacking controls over the program's operation can be established. A continuous chain of uniform practice will

then exist, running from top to bottom of the organization.

Any evaluation system catering at once to the purposes of employee counseling and management information involves a delicate balance between the needs of the individual and the needs of the organization. The successful operation of these systems in particular require clear understanding on the part of the raters and the kind of control that can best be exercised by review authorities.

Adoption of the above recommendations would undoubtedly demand a net increase in the resources devoted to performance evaluation. In view of the expenditures presently incurred by the programs and the reduction of which does not appear very likely, it is believed that the comparatively small additional outlay would be more than compensated by the improved results.

BIBLIOGRAPHY

BOOKS

Belsley, G. Lyle. Federal Personnel Management and the Transition. Washington: Public Administration Clearing House, 1952.

Fleishman, Edwin A. Studies in Personnel and Industrial Psychology. Homewood, Illinois: Dorsey Press, 1961.

Nigro, Felix A. Public Personnel Administration. New York: Henry Holt and Company, 1959.

Stahl, O. Glen. Public Personnel Administration. New York: Harper and Row, 1962.

Van Riper, Paul. History of the United States Civil Service. White Plains: Row, Peterson, 1958.

Whisler, Thomas L., and Shirley F. Harper. Performance Appraisal. New York: Holt, Rinehart and Winston. 1962.

PERIODICALS

Bayroff, A. G., Helen R. Haggerty, and E. A. Rundquist. "Validity of Ratings as Related to Rating Techniques and Conditions," Personnel Psychology, VII (Spring, 1954), 93-113.

Best, Wallace H. "Some New Directions in Personnel Appraisal," Personnel, XXXIV (September, 1957), 45-50.

Bittner, Reign. "Developing an Employee Merit Rating Procedure," Personnel Psychology, I (1948), 403-432.

Cozan, Lee W. "Forced Choice: Better Than Other Rating Methods?" Personnel, XXXVI (May-June, 1955), 80-83.

Fiedler, Fred E. "Assumed Similarity Measures as Predictors of Team Effectiveness," Journal of Abnormal and Social Psychology, XXXIX (1954), 381-388.

Flanagan, John C., and Robert K. Burns. "The Employee Performance Record," Harvard Business Review, (September-October, 1957), 95-102.

Glickman, A. S. "Effects of Negatively Skewed Ratings on Motivations of the Rated," Personnel Psychology, VIII (Spring, 1955), 39-47.

Kellogg, M. S. "Appraising the Performance of Management Personnel," Personnel, XXXI (March, 1955), 442-455.

Mahler, Walter R., and Guyot Frazier. "Appraisal of Executive Performance," Personnel, XXXI (March, 1955), 429-441.

Maloney, P. W., and J. R. Hinrichs. "A New Tool for Supervisory Self-development," Personnel, (July-August, 1959), 46-53.

Mitchell, James M. "Recent Progress in Federal Personnel Administration," Public Personnel Review, (October, 1950) 180-184.

"Personnel Opinions," Public Personnel Review, XXIII (July, 1962), 201-205.

Richards, Kenneth E. "Some New Insights Into Performance Appraisal," Personnel, XXXVII (July, 1960), 28-38.

Stockford, Lee, and H. W. Bissel. "Factors Involved in Establishing a Merit-Rating Scale," Personnel, XXVI (September, 1949), 94-116.

Taylor, Erwin K., and Roy Hastman. "Relation of Format and Administration to the Characteristics of Graphic Rating Scales," Personnel Psychology, IX (Summer, 1956), 181-206.

Wherry, Robert J., and Douglas H. Fryer. "Buddy Ratings: Popularity Contest or Leadership Criteria?" Personnel Psychology, II (Summer, 1949), 147-160.

OTHER PUBLICATIONS

Commission on Organization of the Executive Branch of the Government. Personnel and Civil Service. Washington: Government Printing Office, 1955. 50-65.

Commission on Organization of the Executive Branch of the Government. Task Force Report on Personnel and Civil Service. Washington: Government Printing Office, 1955. 238-241.

Commission on Organization of the Executive Branch of the Government. Personnel Management. Washington: Government Printing Office, 1949. 40.

Public Law 873, Chapter 1123. 81st Congress, 2d Session. The text of the act also may be found in the Federal Personnel Manual, Chapter Z1.

Senate Subcommittee on Federal Manpower Policies. Supervisory Selection in the Federal Government. Washington: Government Printing Office, 1952. 7.

The American Assembly. The Federal Government Service. New York: Columbia University Press, 1954.

Sayre, Wallace S. "The Public Service," Goals for Americans. New York: Prentice Hall, 1963. 285-296.

thes1415

Better performance appraisal for a better



3 2768 002 03482 9

DUDLEY KNOX LIBRARY